


## RESEARCH ARTICLE

# Nonlinear manifold learning in functional magnetic resonance imaging uncovers a low-dimensional space of brain dynamics

Siyuan Gao<sup>1</sup> | Gal Mishne<sup>2,3</sup> | Dustin Scheinost<sup>1,4,5,6</sup> 

<sup>1</sup>Department of Biomedical Engineering, Yale University, New Haven, Connecticut

<sup>2</sup>Hacıoğlu Data Science Institute, University of California San Diego, La Jolla, California

<sup>3</sup>Neurosciences Graduate Program, University of California San Diego, La Jolla, California

<sup>4</sup>Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, Connecticut

<sup>5</sup>Department of Statistics and Data Science, Yale University, New Haven, Connecticut

<sup>6</sup>Child Study Center, Yale School of Medicine, New Haven, Connecticut

## Correspondence

Dustin Scheinost, Department of Biomedical Engineering, Yale University, New Haven, CT 06520, USA.

Email: [dustin.scheinost@yale.edu](mailto:dustin.scheinost@yale.edu)

## Funding information

NIH Roadmap for Medical Research, Grant/Award Numbers: PL1NS062410, PL1MH083271, RL1LM009833, RL1MH083270, RL1DA024853, RL1MH083269, RL1MH083268, UL1-DE019580; the NIH Blueprint for Neuroscience Research, Grant/Award Number: 1U54MH091657; National Institutes of Health, Grant/Award Number: MH121095

## Abstract

Large-scale brain dynamics are believed to lie in a latent, low-dimensional space. Typically, the embeddings of brain scans are derived independently from different cognitive tasks or resting-state data, ignoring a potentially large—and shared—portion of this space. Here, we establish that a shared, robust, and interpretable low-dimensional space of brain dynamics can be recovered from a rich repertoire of task-based functional magnetic resonance imaging (fMRI) data. This occurs when relying on nonlinear approaches as opposed to traditional linear methods. The embedding maintains proper temporal progression of the tasks, revealing brain states and the dynamics of network integration. We demonstrate that resting-state data embeds fully onto the same task embedding, indicating similar brain states are present in both task and resting-state data. Our findings suggest analysis of fMRI data from multiple cognitive tasks in a low-dimensional space is possible and desirable.

## KEYWORDS

diffusion maps, dynamic connectivity, integration, participation coefficient, segregation

## 1 | INTRODUCTION

Understanding large-scale brain dynamics is a major goal of modern neuroscience (Jorgenson et al., 2015). However, due to the high-dimensional nature of brain patterns, how to best operationalize and tackle this problem remains an open question. Nevertheless, the temporal dimensions that explain the observed dynamics are small compared with the number of time points (Cunningham & Byron, 2014). Thus, there is growing evidence to suggest that a low-dimensional space—hidden from direct observation, learned from the data, and derived from many brain regions—may be a

suitable model for studying temporal brain dynamics (Gao & Ganguli, 2015).

These low-dimensional spaces have been observed using a variety of neural recordings and animal models (Ahrens et al., 2012; Churchland et al., 2012; Kobak et al., 2016; Mishne et al., 2016; Santhanam et al., 2009). Research suggests that linear methods, such as principal component analysis (PCA), are appropriate when recorded temporal data comes from simple stimuli that project onto a limited area within a manifold (Cunningham & Byron, 2014). However, data from richer tasks often project onto a larger portion of the manifold, violating linear approximations (Cunningham & Byron, 2014; Gallego,

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

Perich, Miller, & Solla, 2017). Nonlinear dimensionality reduction methods, like diffusion maps (Coifman & Lafon, 2006), can overcome this limitation by integrating local similarities into a global representation, which had better reflect the underlying temporal dynamics in neural recordings.

Similar concepts have emerged in human functional magnetic resonance imaging (fMRI) studies to quantify moment-to-moment changes in activity and connectivity (Hutchison et al., 2013; Preti, Bolton, & Van De Ville, 2017). As with related research on temporal recordings from animal models, dimensionality reduction methods are used to project the fMRI time series onto a low-dimensional space (Allen et al., 2014; Monti et al., 2017; Shine et al., 2016; Shine et al., 2019). From the low-dimensional space, characteristic brain states—or distinct, repeatable patterns of brain activity—are used to quantify brain dynamics. Predominantly, these studies have relied on linear methods (Allen et al., 2014; Monti et al., 2017; Shine et al., 2016; Shine et al., 2019). However, given the rich repertoire of tasks available in human fMRI, a manifold derived from nonlinear methods may better capture the underlying geometry of the low-dimensional space.

To address this, we use a recently introduced extension of diffusion maps, 2-step Diffusion Maps (2sDM; Gao, Mishne, & Scheinost, 2019). 2sDM extracts common variability between individuals by performing dimensionality reduction of a third-order tensor in a two-stage manner. In the first stage, time series data from each individual are embedded into a low-dimensional Euclidean space. In the second stage, embedding coordinates for the same time point from different individuals are concatenated for use in a second embedding. The second stage embeds similar time points across subjects to obtain a low-dimensional group-wise representation of those time points. This two-stage approach avoids directly comparing brain activation across subjects, which can be imprecise without proper alignment (Haxby et al., 2011). While diffusion maps have been applied to fMRI data (Margulies et al., 2016; Nenning et al., 2020), we aim to embed the time dimension rather than the spatial dimension.

We used 2sDM to embed time series from a rich repertoire of tasks onto a single low-dimensional manifold in two fMRI datasets: the Human Connectome Project and the UCLA Consortium for Neuropsychiatric Phenomics. By using multiple tasks spanning a range of cognitive functions and loads, we obtain a more even sampling of the original high-dimensional space of recurring patterns of brain dynamics (Cunningham & Byron, 2014; Gallego et al., 2017) to better project individual time points onto a low-dimensional manifold. Our embedding positioned different tasks by their cognitive load. Thus, it enables scans to be compared both within the same task and across different tasks. As our embedding also has a clear clustering structure, downstream analyses that are based on brain states or low-dimensional trajectories are also straightforward to perform based on the embedding. Additionally, we embedded resting-state data into the same task embedding to investigate differences in brain dynamics between resting-state and task performance. These results suggest that manifold learning can uncover an interpretable low-dimensional embedding for the study of brain dynamics in fMRI data.

## 2 | METHODS

### 2.1 | Dataset and imaging parameters

Data was obtained from the Human Connectome Project (HCP) 900 Subject release (Van Essen et al., 2013). We use fMRI data collected while 390 participants performed six tasks (gambling, motor, relational, social, working memory—WM, and emotion). We restrict our analyses to those subjects who participated in all nine fMRI conditions (seven task, two rest), whose mean frame-to-frame displacement is less than 0.1 mm and whose maximum frame-to-frame displacement is less than 0.15 mm, and for whom the task block order is the same as other subjects ( $n = 390$ ). All fMRI data were acquired on a 3 T Siemens Skyra using a slice-accelerated, multiband, gradient-eco, echo planar imaging (EPI) sequence (TR = 720 ms, TE = 33.1 ms, flip angle = 52°, resolution = 2.0 mm<sup>3</sup>, multiband factor = 8). Images acquired for each subject include a structural scan and 18 fMRI scans (working memory [WM] task, incentive processing [gambling] task, motor task, language processing task, social cognition task, relational processing task, emotion processing task, and two resting-state scans; two runs per condition [one LR phase encoding and one RL phase encoding run]) split between two sessions.

The UCLA Consortium for Neuropsychiatric Phenomics (CNP; Poldrack et al., 2016) dataset is used for replication. Similar to the standards for the HCP dataset, we restrict our analyses to those subjects who participated in all five fMRI conditions (four task, one rest), whose mean frame-to-frame displacement is less than 0.1 mm and whose maximum frame-to-frame displacement is less than 0.15 mm. Seventy-seven healthy controls are retained. These participants performed four tasks (paired memory retrieval task—PAMRET, paired memory encoding task—PAMENC, spatial working memory task—SCAP, task switching task—TASKSWITCH). Details of the image acquisition parameters have been published elsewhere (Poldrack et al., 2016). In brief, all data were acquired on one of two 3T Siemens Trio scanners at UCLA. Functional MRI data were collected using a T2\*-weighted EPI sequence with the following parameters: slice thickness = 4 mm, 34 slices, TR = 2 s, TE = 30 ms, flip angle = 90°, matrix 64 × 64, FOV = 192 mm, oblique slice orientation. Images acquired for each subject include a structural scan and seven fMRI scans (balloon analog risk task (BART), paired-associate memory retrieval (PAMRET), paired-associate memory encoding (PAMENC), spatial capacity task (SCAP), stop signal task (SST), task-switching task (TASKSWITCH) and breath holding task).

As 2sDM requires time series to be synchronized across individuals (i.e., different individuals encounter the same task condition at the same time point), the language task from the HCP and the stop signal task, balloon analog risk task, and breath hold task from the CNP were not included. These tasks are self-paced. Participants finished blocks at different times, causing the task block to be unsynchronized across participants.

### 2.2 | fMRI processing

For the HCP dataset, the HCP minimal preprocessing pipeline was used (Glasser et al., 2013), which includes artifact removal, motion

correction, and registration to standard space. For the CNP dataset, structural scans were skull-stripped using OptiBet (Lutkenhoff et al., 2014) and registered to the MNI template using a validated algorithm in BiImage Suite (Joshi et al., 2011; Scheinost et al., 2017). Slice time and motion correction were performed in SPM8. For both datasets, all subsequent preprocessing was performed using image analysis tools available in BiImage Suite and included standard preprocessing procedures (Finn et al., 2015). Several covariates of no interest were regressed from the data including linear and quadratic drifts, mean cerebral-spinal-fluid signal, mean white-matter signal, and mean gray matter signal. For additional control of possible motion related confounds, a 24-parameter motion model (including six rigid-body motion parameters, six temporal derivatives, and these terms squared) was regressed from the data. The data were temporally smoothed with a Gaussian filter (approximate cutoff frequency = 0.12 Hz). Mean frame-to-frame displacement yielded seven motion values per subject, which were used for subject exclusion and motion analyses. We restricted our analyses to subjects whose maximum frame-to-frame displacement was less than 0.15 mm and mean frame-to-frame displacement was less than 0.1 mm. This conservative threshold for exclusion due to motion was used to mitigate the effect of motion on the embedding. We used the Shen 268-node atlas to extract time series from the fMRI data for further analysis (Shen, Tokoglu, Papademetris, & Constable, 2013). Time series used for the embedding were the average of the basis of the “raw” task time courses, with no removal of task-evoked activity, for each node in the atlas. Finally, 2sDM was applied to embed a third-order tensor of fMRI data (individual  $\times$  region  $\times$  time) onto a single low-dimensional manifold.

### 2.3 | 2-step diffusion maps (2sDM)

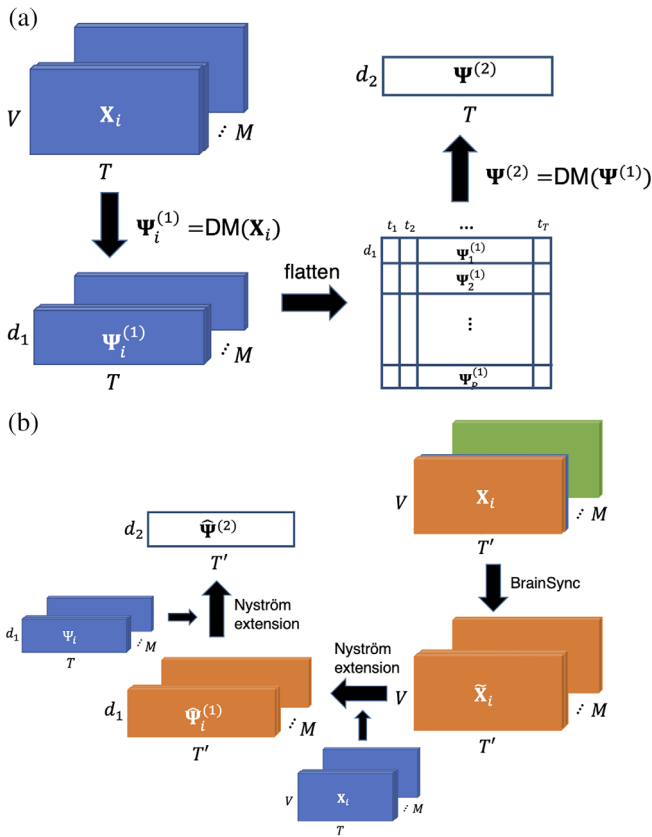
Diffusion maps are part of a broad class of manifold learning algorithms. Specifically, diffusion maps provide a global description of the data by considering only local similarities and are robust to noise perturbations. The new nonlinear representation provided by diffusion maps reveals underlying intrinsic parameters governing the data (Coifman & Lafon, 2006). We briefly describe the diffusion maps algorithm in general and in the following its application to fMRI data as used in our approach. Given a dataset of  $n$  points  $\{\mathbf{x}_i\}_{i=1}^n$ , a pairwise similarity matrix  $\mathbf{S}$  between pairs of data points  $\mathbf{x}_i$  and  $\mathbf{x}_j$  is constructed, for example using the Gaussian kernel  $w_\epsilon(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/\epsilon)$ , where  $\|\cdot\|$  denotes the Euclidean norm. The Gaussian kernel omits global information, resulting in non-linearity. Then the rows of the similarity matrix are normalized by  $\mathbf{P} = \mathbf{D}^{-1}\mathbf{S}$ , where  $\mathbf{D}_{ii} = \sum_j \mathbf{S}_{ij}$  is the degree of point  $\mathbf{x}_i$ . This creates a random walk matrix on the data with entries set to  $p(\mathbf{x}_i, \mathbf{x}_j) = w_\epsilon(\mathbf{x}_i, \mathbf{x}_j)/d(\mathbf{x}_i)$ , where  $d(\mathbf{x}_i) = \mathbf{D}_{ii}$ , the degree of point  $\mathbf{x}_i$ . Taking the  $t$ -th powers of the matrix  $\mathbf{P}$  is equivalent to running the Markov chain corresponding to the random walk on the data forward  $t$  times. The corresponding kernel  $p_t(\cdot, \cdot)$  can then be interpreted as the transition probability between two points in  $t$  time steps. The matrix  $\mathbf{P}$  has a complete sequence of bi-orthogonal left and right eigenvectors  $\phi_i$  and  $\psi_i$ , respectively, and a corresponding sequence of eigenvalues  $1 = \lambda_0 \geq |\lambda_1| \geq |\lambda_2| \geq \dots$ .

Diffusion maps are a nonlinear embedding of the data points into a low-dimensional space, where the mapping of point  $\mathbf{x}$  is defined as  $\Psi(\mathbf{x}) = (\lambda_1^t \psi_1(\mathbf{x}), \lambda_2^t \psi_2(\mathbf{x}), \dots, \lambda_k^t \psi_k(\mathbf{x}))$ , where  $t$  is the diffusion time. Note that  $\psi_0$  is neglected because it is a constant vector. The diffusion distance  $D_t^2(\mathbf{x}, \mathbf{y})$  between two data points is defined as:

$$D_t^2(\mathbf{x}, \mathbf{y}) = \sum_z \frac{(p_t(\mathbf{x}, \mathbf{z}) - p_t(\mathbf{y}, \mathbf{z}))^2}{\phi_0(\mathbf{z})} \quad (1)$$

where  $\phi_0$  represents the stationary distribution of the random walk described by the random walk matrix  $\mathbf{P}$ . This measures the similarity of two points by the evolution in the Markov chain and the distance characterizes the probability of transition from  $\mathbf{x}$  or  $\mathbf{y}$  to the same  $\mathbf{z}$  point in  $t$  time steps. Two points are closer with smaller  $D_t^2(\mathbf{x}, \mathbf{y})$  if there is a large probability of transition from  $\mathbf{x}$  to  $\mathbf{y}$  or vice versa, suggesting that there are more short paths connecting them. It is thus robust to noise as it considers all the possible paths between two points and is thus less sensitive to noisy connections. It was proved that the  $k$ -dimensional diffusion maps  $\Psi$  embed data points into a Euclidean space  $\mathbb{R}^k$  where the Euclidean distance approximates the diffusion distance (Coifman & Lafon, 2006). In practice, eigenvalues of  $\mathbf{P}$  typically exhibit a spectral gap such that the first few eigenvalues are close to one with all additional eigenvalues being much smaller than one. Thus, the diffusion distance can be well approximated by only the first few eigenvectors. Therefore, we can obtain a low-dimensional representation of the data by considering only the first few eigenvectors of the diffusion maps. Intuitively, diffusion maps embed data points closer when it is harder for the points to escape their local neighborhood within time  $t$ .

To obtain a groupwise low-dimensional representation of dynamics, a hierarchical diffusion maps-based manifold learning framework, 2-step diffusion maps (2sDM; Gao et al., 2019), was used to reduce the dimensionality of high-dimensional multi-individual fMRI time series. The framework is illustrated in Figure 1a. Under the assumption that individuals' fMRI responses are time-synchronized, the fMRI BOLD time series data are organized as three-dimensional array  $\mathbf{X} \in \mathbb{R}^{M \times V \times T}$  (number of individuals  $M$ , number of regions or voxels  $V$ , number of time points  $T$ ). In the first step of 2sDM, each individual is processed separately, by applying diffusion maps to the fMRI time series of every single individual  $\mathbf{X}_{i\cdot} \in \mathbb{R}^{V \times T}$ , thereby obtaining a  $d_1$ -dimensional temporal embedding of each individual  $\Psi_i^{(1)} \in \mathbb{R}^{d_1 \times T}$ . Then, in the second step, we first concatenate the new representations from all individuals into a matrix  $\Psi^{(1)} \in \mathbb{R}^{(Md_1) \times T}$ , such that each time-point is now represented by the embeddings of that time-frame aggregated from all  $M$  subjects. Then, a second-round diffusion embedding is performed, further reducing the dimensionality of every time-frame to  $d_2$  and the final time-frame representation with multi-individual similarity is  $\Psi^{(2)} \in \mathbb{R}^{d_2 \times T}$ , where  $d_1$  and  $d_2$  are predetermined parameters that are smaller than  $V$ . The concatenation and two-round embeddings are theoretically supported by the theorem that low-dimensional diffusion maps approximate the diffusion distance between time-frames (Gao et al., 2019): The distance between two frames  $\Psi^{(2)}(t_1)$  and  $\Psi^{(2)}(t_2)$  approximates the average diffusion



**FIGURE 1** Schematic of manifold learning framework. (a) 2sDM algorithm framework for time-synchronized multi-individual fMRI time series. (b) 2-step out-of-sample extension framework with BrainSync for new fMRI time points. Mathematical notations in the figure are the same as those used in the corresponding Section 2. fMRI, functional magnetic resonance imaging; 2sDM, 2-step Diffusion Maps

distance between those time-frames across all individuals. We used  $d_1 = 7$  and  $d_2 = 3$  in our experiment. It is worth noting that the embedding results were robust under a certain range of different  $d_1$  and  $d_2$  (related discussion in Data S1 and Figure S3).

To reveal the progression of brain dynamics during tasks, we calculated temporal trajectories (Shine et al., 2019) for each task block by connecting points in the embedding in a temporal order. As the tasks involve the same task blocks with repetitions (i.e., WM task consists of interleaved blocks of 0-back and 2-back with the same length), we averaged the time-frames belonging to the same task block to obtain a smoothed representative trajectory of each task. Time frames corresponding to the cue or fixation between tasks blocks were not included.

To summarize the embedding into a more compact and easier to analyze structure, we performed  $k$ -means clustering based on the first three embedding dimensions to cluster time points sharing similar brain activation patterns into discrete brain states. The Calinski-Harabasz criterion (ratio between the within-cluster dispersion and the between-cluster dispersion) was used to determine the number of clusters, evaluating values of  $k = \{2, \dots, 10\}$  (Caliński & Harabasz, 1974).

To illustrate that nonlinear discovers structure that linear methods cannot recover, we used a 2-step PCA framework, similar to 2sDM. In the first step, a separate PCA is applied to each individual's fMRI time series  $X_{i,\dots} \in \mathbb{R}^{V \times T}$ , resulting into a  $d_1$ -dimensional linear temporal embedding of each individual  $c_i^{(1)} \in \mathbb{R}^{d_1 \times T}$ , where the first  $d_1$  principal components with the maximum variances are included. Then each individual's embedding is concatenated along the time dimension form to a matrix  $C^{(1)} = [c_1^{(1)}, c_2^{(1)}, c_3^{(1)}, \dots, c_M^{(1)}] \in \mathbb{R}^{(Md_1) \times T}$ . A second-round PCA is performed to further reduce the dimensionality of this concatenated matrix. Each time frame is then embedded into  $d_2$  dimensions and the final time-frame representation with multi-individual similarity is  $C^{(2)} \in \mathbb{R}^{d_2 \times T}$ . Same as 2sDM, we used  $d_1 = 7$  and  $d_2 = 3$  in our experiment.

## 2.4 | Dynamic connectivity

To relate our task embedding to previously used handcrafted features (Shine et al., 2016), we calculated the participation coefficient  $B_T$  using sliding-window-based functional connectivity and then averaged  $B_T$  across all subjects, as described in previous literature (Shine et al., 2016). In this article, handcrafted features refer to features that are designed manually, such as  $B_T$  that is used here to characterize the integration and segregation pattern of the brain network. The dynamic functional connectivity is calculated by the multiplication of temporal derivatives (MTD; Shine et al., 2015). MTD is calculated as the point-wise product of the temporal derivatives of paired nodes ( $i, j$ )'s time series:

$$MTD_{ij,t} = \frac{1}{w} \sum_{k=t}^{k=t+w} \frac{\Delta x_{i,k} \times \Delta x_{j,k}}{\sigma_{\Delta x_i} \times \sigma_{\Delta x_j}} \quad (2)$$

where  $\Delta x_{i,t} = x_{i,t} - x_{i,t-1}$  is the temporal derivative of node  $i$  at time  $t$  with time series  $x$ ,  $\sigma_{\Delta x_i}$  is the SD of the  $\Delta x_i$  over the entire time course and  $w$  is the window length. At each time point, the dynamic functional connectivity is calculated as the averaged MTD over a sliding time window in order to reduce high-frequency noise. We chose the length of the sliding window  $w = 14$  for HCP ( $\sim 10.8$  s) and  $w = 4$  for CNP ( $\sim 10$  s), based on previous literature (Shine et al., 2016).

The participation coefficient  $B_T$  characterizes the extent to which a region connects across all modules, where modules are normally defined a priori from community detection methods that identify a set of nodes as a module that are more strongly connected to each other than nodes from another set. The time-resolved community structure was used here according to (Shine et al., 2016) and it was estimated by the Louvain modularity algorithm from the Brain Connectivity Toolbox (Rubinov & Sporns, 2010). The participation coefficient for a region  $i$  at time  $T$  is calculated as:

$$B_{i,T} = 1 - \sum_{s=1}^{N_M} \left( \frac{k_{i,s,T}}{k_{i,T}} \right)^2 \quad (3)$$

where  $k_{i,s,T}$  is the strength of the positive connections of node  $i$  to nodes in module  $s$  at time  $T$ ,  $k_{i,T}$  is the sum of strengths of all positive

connections of node  $i$  at time  $T$  and  $N_M$  is the number of modules, or time-resolved communities in our setting. The participation coefficient of a region is therefore close to 1 if its links are uniformly distributed among all the modules and 0 if all its links are within its own module. The whole brain participation coefficient  $B_T$  represents the average of  $B_{i,T}$  from each region and thus represents the integration and segregation pattern of the brain.  $B_T$  is closer to 1 if our whole brain is more integrated and closer to 0 if our whole brain is more d.

## 2.5 | 2-Step out-of-sample extension (OOSE) for resting-state fMRI

To investigate the generalization of the task manifold and associated brain states, resting-state data were embedding onto the manifold. One of the challenges in nonlinear dimensionality reduction is to extend new data points to the embedding space. Unlike linear dimensionality reduction methods like PCA, there is no explicit mapping from the original features to the new coordinates. Moreover, appending the new data and redoing the dimensionality reduction is often computationally costly. To deal with this, we specially designed a corresponding 2-step out-of-sample extension (OOSE) framework to embed new time points onto the existing temporal manifold.

The framework is illustrated in Figure 1b. The framework follows a similar two-step structure to 2sDM where the Nyström extension (Fowlkes, Belongie, Chung, & Malik, 2004) (a nonparametric OOSE method, details provided in Data S1) is used to approximate the reduced representation of the new time series in each step. Specifically, given new fMRI time series  $\tilde{X}_{i,\dots} \in \mathbb{R}^{V \times T}$ ,  $i = 1, \dots, M$  from the same group of individuals used for 2sDM embedding, we first approximate the eigenvectors  $\hat{\Psi}_i^{(1)}$  for each individual using Nyström extension. Then we concatenate all the individuals' eigenvectors  $\hat{\Psi}_i^{(1)}$  as the new data points and approximate its eigenvectors  $\hat{\Psi}^{(2)}$  as the final representation.

As the 2sDM algorithm requires the task designs across individuals to be the same, this prevents embedding multi-individual resting-state fMRI time series directly, which is also a problem for any other scans that are not time-synchronized, for example, the language task in the HCP dataset. To synchronize different individual's time series, we used BrainSync, a framework that synchronizes fMRI time series across individuals (Joshi, Chong, Li, Choi, & Leahy, 2018). BrainSync synchronizes one individual's time series data  $\mathbf{Y} \in \mathbb{R}^{V \times T}$  to another reference individual  $\mathbf{X} \in \mathbb{R}^{V \times T}$  by finding an optimal orthogonal transformation that minimizes summed moment-to-moment squared error  $\mathbf{O}^s = \arg \min_{\mathbf{O} \in \mathbf{O}(T)} \|\mathbf{X} - \mathbf{Y}\mathbf{O}^t\|^2$ . The validity of BrainSync is based on the observation that the resting-state fMRI data exhibit similar connectivity patterns across individuals and thus an orthogonal transformation is able to align two individual's time points as the time points lie on a hypersphere. The problem can be solved by the Kabsch algorithm (Kabsch, 1976). The  $T \times T$  cross-correlation matrix  $\mathbf{X}^t \mathbf{Y}$  is first formed and its singular value decomposition can be calculated as  $\mathbf{X}^t \mathbf{Y} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^t$ . The optimal  $\mathbf{O}^s$  can be found by  $\mathbf{O}^s = \mathbf{U} \mathbf{V}^t$  and  $\mathbf{Y}$  can be synchronized to  $\mathbf{X}$  by  $\mathbf{Y}\mathbf{O}^{st}$ . By choosing a random individual as the

reference, BrainSync was applied to all the other individuals and their time series were synchronized to the reference individual. After synchronizing across individuals, we then used the 2-step OOSE framework to extend them onto the task manifold and find the temporal representation of resting-state fMRI for the reference individual.

To validate the 2-step OOSE framework, we used the task fMRI data to cross-validate the accuracy of the OOSE framework. Using leave-one-task-out cross-validation, a single task was held out when generating the 2sDM manifold. The left-out task was then embedded in the new manifold using our OOSE framework and compared with the original embedding created using all tasks. If the held-out task's extended coordinates are similar to the coordinates of the original embedding, it suggests that the OOSE framework is accurate.

## 2.6 | Characterizing changes in brain states

By utilizing the temporal order of time points, we characterized the brain dynamics across the four brain states by state transition probability and dwell time. State transition probabilities were calculated based on the temporally adjacent time points' brain states. From these state transition probabilities, a stochastic matrix and the dwelling times (i.e., the stationary probability distribution of the stochastic matrix) were calculated and visualized as Markov chain models. The stationary distribution of the Markov transition matrix  $\mathbf{P}_{\text{trans}}$  is defined as the distribution that does not change under application of the transition matrix  $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}_{\text{trans}}$ , which is the left eigenvector of  $\mathbf{P}_{\text{trans}}$ . It represents the distribution to which the Markov process converges. It was used in our experiment to represent the dwell-time distribution of discrete brain states. As tasks putatively put a participant into certain states (as opposed to the unconstrained nature of the resting state), we investigated differences in the temporal dynamics of state switching during task and rest. We calculated entropy—a measure of the randomness—of the transition probability from one brain state to the other states. Entropy of a discrete probability distribution measures the uncertainty of the outcome. It is calculated as the negative expectation of the logarithm of the probability mass function's value  $S = -\sum_i P_i \log P_i = -E_p[\log P]$ . In our experiment, entropy of the brain state transition probability was used to assess the randomness of brain state transitioning with lower entropy representing more easy-to-predict brain state transition dynamics. Greater entropy indicates a less predictable transition from one state to another.

## 2.7 | Experimental design and statistical analysis

No statistical methods were used to predetermine sample sizes. Other than the stated exclusion criteria for motion and complete imaging data, no participants and data points were excluded from the analysis. Following exclusion for motion, there was no significant correlation between motion and the embedding dimension. Parametric statistics (e.g.,  $t$ -test, correlation, and chi-squared test) were used when appropriate.



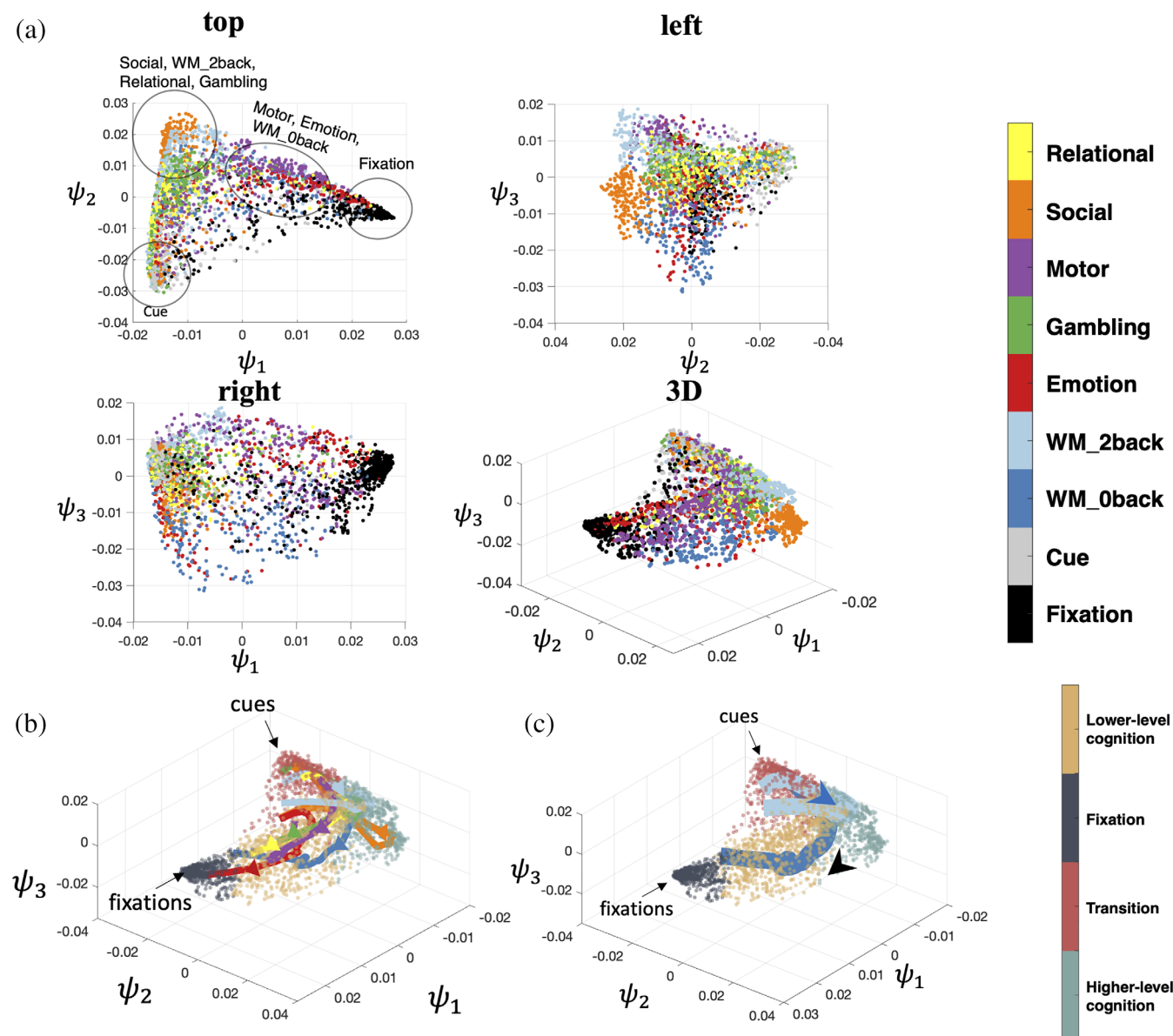
## 2.8 | Data availability

The HCP data used in this study are publicly available from the ConnectomeDB database (<https://db.humanconnectome.org>). The CNP data used in this study are publicly available from OpenNeuro.org (<https://openneuro.org/datasets/ds000030>). MATLAB scripts to run the 2sDM analyses can be found at (<https://github.com/carricky/2sDM>). Biolmage Suite tools used for analysis can be accessed at (<https://bioimagesuiteweb.github.io/>).

## 3 | RESULTS

### 3.1 | Brain dynamics during tasks embed onto a low-dimensional space

Although each task is different in many ways, individual time points in the fMRI data from all tasks mapped onto a single low-dimensional manifold (Figure 2a). We find a global representation across multiple tasks that positioned tasks with similar cognitive loads together. By



**FIGURE 2** Nonlinear embedding of fMRI time series data. (a) 2sDM embedding of seven tasks (relational, social, motor, gambling, emotion, working memory 2-back, and working memory 0-back) from the HCP dataset. Four different views of the manifold are shown. Each point in these subplots represents a single time point and is colored by the task type. (b) Averaged temporal trajectory of each task with the embedding colored by the corresponding brain state as the background. (c) WM task's 0-back and 2-back task blocks visualized separately with major cues and fixations points annotated. Arrows show the progression direction of the trajectory. Trajectory in (b) and (c) uses the same colormap as (a). fMRI, functional magnetic resonance imaging; HCP, Human Connectome Project; 2sDM, 2-step Diffusion Maps

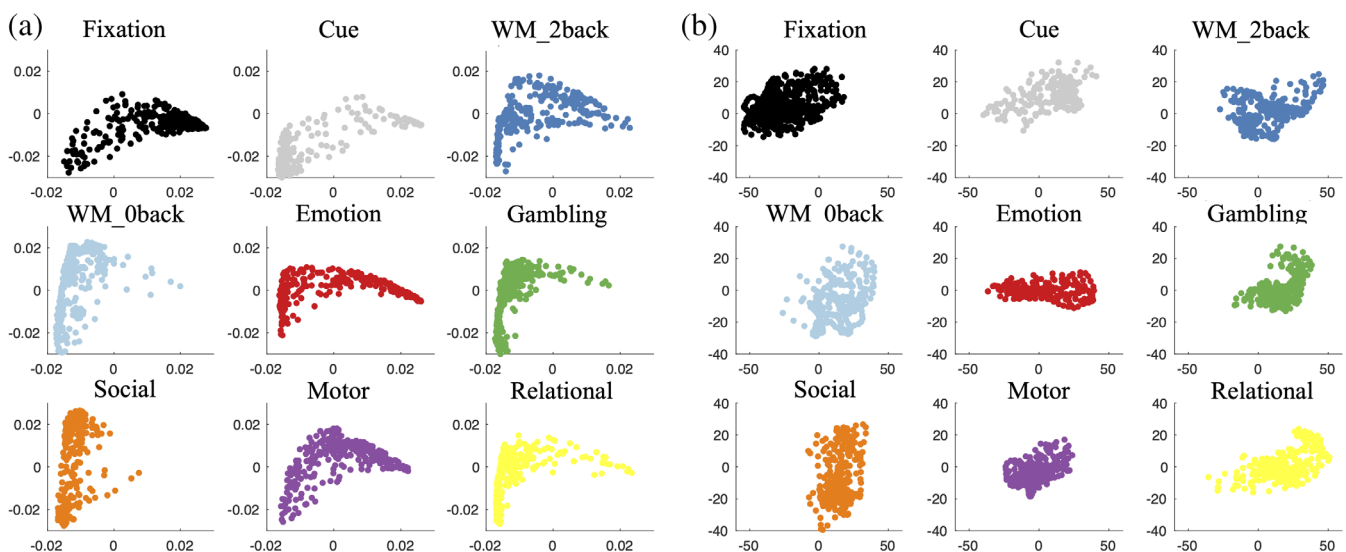
embedding multiple tasks together, rather than in isolation, the closeness of different blocks and tasks in the manifold suggest that similar, recurring patterns of brain dynamics exist across a variety of tasks. For example, in the manifold, the 2-back blocks of the WM task are significantly ( $t=201.9, p < .01, df = 175,102$ ) closer to time points from the gambling task (Euclidean distance:  $0.0258 \pm 0.0096$ ) than the 0-back blocks of the WM task (Euclidean distance:  $0.0355 \pm 0.0100$ ), despite the fact that the 2-back and 0-back blocks were collected in the same fMRI run. The 2-back blocks of the WM task and the gambling task both entail a higher cognitive load. In contrast, the 0-back blocks of WM task overlap with the motor task in the  $\psi_1$  coordinate. These tasks are simpler response tasks and less cognitively demanding. Overall, these time points are positioned based on the similarity of the cognitive load at that time point, instead of by task.

For all tasks, the average trajectories from each task are found to start near the corner where cues (task cues preceding each task block) reside and end in the other corners depending on the task progression patterns. These smooth trajectories indicate that the embedding preserves proper temporal associations between blocks when arranging time points in discrete states. All trajectories start in the same brain state (bottom left corner) and are clustered together at the start. Then the trajectories from each task traverses through different proportions of the embeddings. As can be expected, the paths of these temporal trajectories depend on the cognitive processes of the task block. For example, the 2-back task traverses through the upper part of the manifold (higher value in terms of  $\psi_3$ ), and, in contrast, the 0-back task traverses through the lower part of the manifold (Figure 2b). More detailed analysis of the task trajectories is included in the third subsection of Section 3 (Figure 5c). Moreover, as can be seen from the top 20 eigenvalues of the diffusion matrix the spectrum decays rapidly, which suggests that the data is low-dimensional (Figure S1).

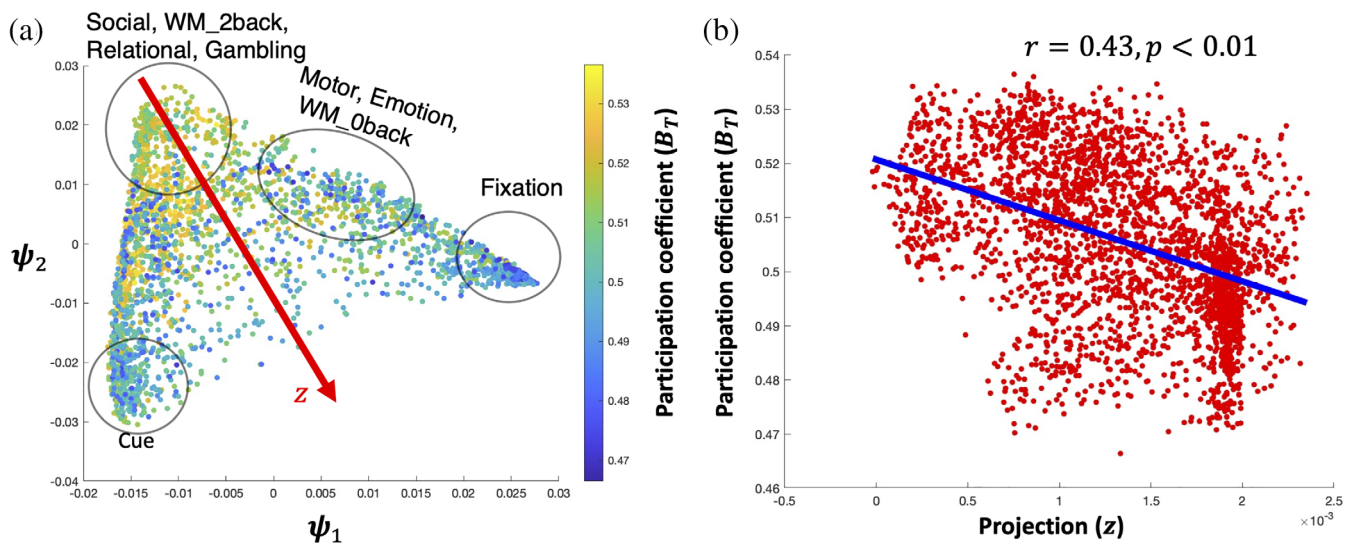
When projecting task fMRI time-frames into 3D space using the first three coordinates of PCA, the embedding separates the different task, rather than the different cognitive components shared across tasks (Figure S4). In the 2sDM embedding, all tasks have a similar arc-shaped and fill a wider range in the embedding space (Figure 3). Individual time points are clustered accordingly to their putative cognitive load. For example, time points in the fixation and cue task block concentrate in one of the two corners of the arc. While time points for more demanding cognitive loads, cluster near the peak of the arc. In contrast, for the 2-step PCA embedding, time points are clearly only clustered by task and no consistent pattern across tasks in the embedding are observed. In other words, in the 2sDM, a clear gradient of cognitive demand that is common across the tasks is observed. No such pattern is observable with the 2-step PCA embedding. Other detailed comparison between 2-step PCA and 2sDM embeddings are included in Figures S4–S7.

### 3.2 | Task embedding captures handcrafted features in an unsupervised manner

In Figure 4, each time point in our task embedding is colored by its subject-averaged  $B_T$ , showing a clear pattern of decreasing  $B_T$  starting from the top left corner of the embedding; higher  $B_T$  at the top of the embedding (i.e., high cognitive load tasks such as social, 2back, relational and gambling) indicates time points of higher integration and lower  $B_T$  at the tails of the embedding (i.e., cues and fixations) indicates time points of higher segregation ( $r(z, B_T) = 0.41, df = 3018, p < .01$ , where  $z$  is the projection coordinates of points onto the diagonal of the triangular embedding; Figure 4b).



**FIGURE 3** Each task independently visualized in the embedding. (a) 2-step Diffusion Maps embedding. (b) 2-step PCA embedding. The x, y axis limits are kept the same within (a) and (b) for better cross-task comparison. The same task block colormap is the same used in Figure 2a. PCA, principal component analysis



**FIGURE 4** 2sDM embedding is related with participation coefficient. (a) 2sDM embedding in HCP dataset colored by the time-resolved  $B_T$ . (b) Scatter plot of the  $B_T$  with the projection onto the diagonal of the embedding structure ( $z$ ). Correlation of  $z$  with  $B_T$  is shown with a line of best fit. Projection direction  $z$  was determined manually as the approximate diagonal direction of the embedding. HCP, Human Connectome Project; 2sDM, 2-step Diffusion Maps

### 3.3 | Operationalizing discrete, recurring brain states from task dynamics

When clustering the task embedding,  $k = 4$  gives the largest Calinski-Harabasz score among a range, suggesting that the embedding has a clear interpretable structure (Figure S6). Based on the task contents of the temporal clusters, we labeled the four brain states as: fixation, transition, lower-level cognition, and higher-level cognition. Functionally reasonable and distinct patterns of activation during the different states are observed, for example, canonical patterns of default mode network activity for the fixation state (Figure 5a). To relate these brain states to previous handcrafted features, we calculated the average  $B_T$  for each brain state (Figure 5b). The four states followed the expected patterns of  $B_T$ , with the higher-level cognition state showing the highest  $B_T$  ( $t = 13.2, p < .01, df = 1527$ ) and the fixation state showing the lowest  $B_T$  ( $t = 14.3, p < .01, df = 1297$ ). The clustering results are similar with an increased number of clusters or of embedding dimensions.

With the help of the four brain states, the dynamic trajectories can further reveal each task's cognitive process (Figure 5c). For example, the motor task's trajectory reveals a dynamic cognitive process as following. In the beginning, the individuals start from the cue state that was the common starting state across the other tasks. Then, the individuals briefly enter the high-cog state, but not deep in the state. Finally, individuals enter and stay in the low-cog state. The trajectory also reveals that, on average, individuals wander towards the fixation state in the middle of the task block, suggesting a fatigue or practice effect. Towards the end of the task block, individuals return deep into the low-cog state and moved towards the cue state for the next task block to start.

Even for tasks like relational and social tasks that both require a certain level of high-level cognitive ability (Shine et al., 2016), there

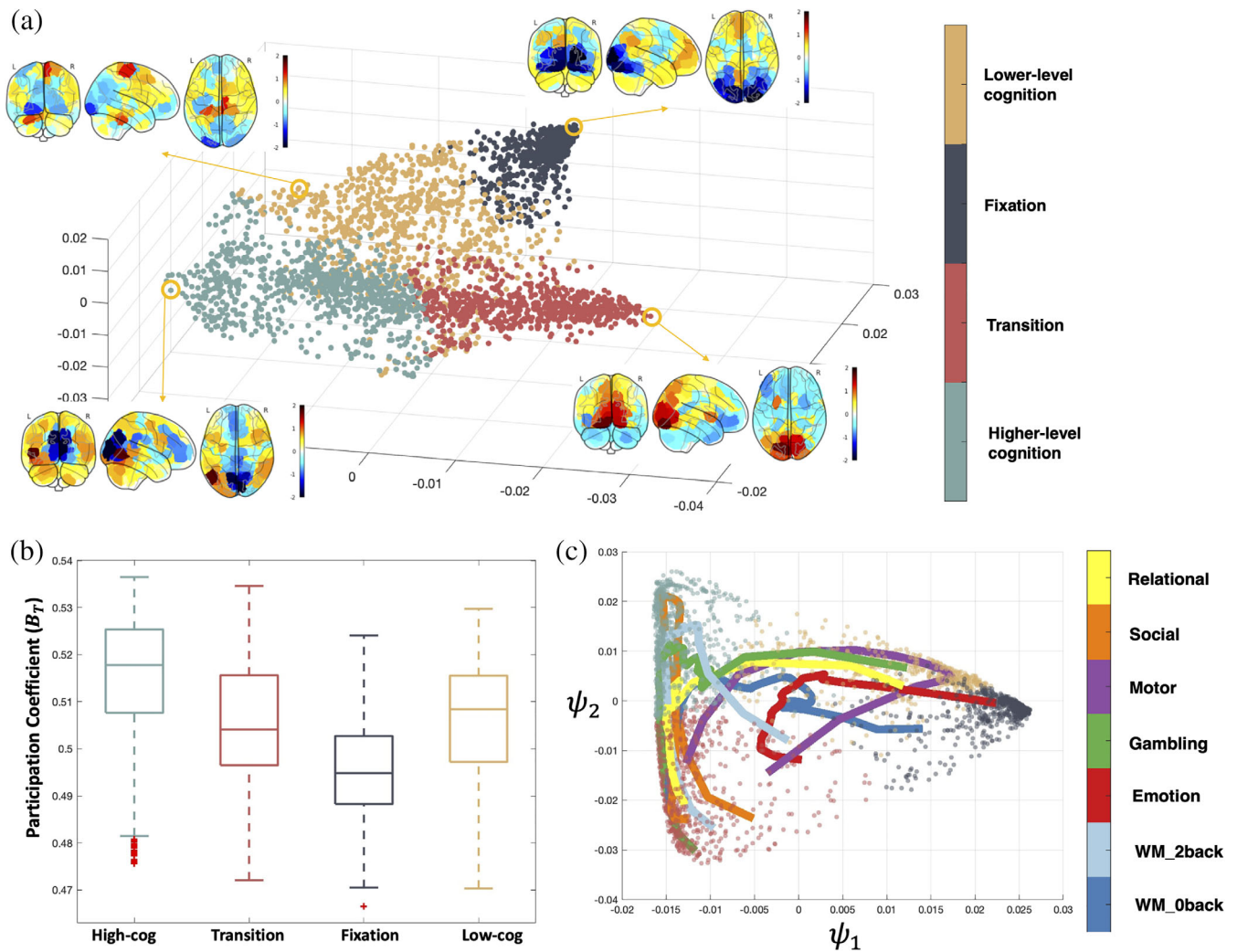
are differences that can be revealed by the trajectories. The relational task starts from the transition cluster, then entered the higher-level cognition cluster and ends in the low-cog state, which suggests a lack of high-level cognitive ability involvement (adaptive to the task design) in the later stage of the relational task blocks. In comparison, the social task starts near the transition cluster, goes deep into the high-cog state and returns to the transition state near the end of the task, which suggests a constant requirement of higher-level cognitive ability. This trajectory view of each task enables a better understanding of the cognitive process and can also help in the future task designs.

The transitions between states are similar for all tasks except for the motor task (which had a high probability of transiting into the lower-level cognition state and out of the higher-level cognition state; Figure 6a). Except for the WM task, which contains an equal proportion of high (2-back) and low (0-back) cognitive loads, dwell times for the four states exhibit a nonuniform distribution ( $\chi^2 > 16.3, d.f. = 3, p < 0.001$ ; Figure 6b), indicating participants spent most of their time in certain limited states in a task-specific manner. For example, the lower-level cognition state occurs most frequently in the motor task, while the higher-level cognitive state dominates in social task time points.

### 3.4 | Brain dynamics during rest embed onto the same recurring brain states which appeared during tasks

Once embedded onto the task manifold, time points from the resting-state data spread across the whole manifold, including parts of the manifold corresponding to higher cognitive loads (Figure 7a). To





**FIGURE 5** Brain states during tasks. (a) Results of  $k$ -means clustering of the task manifold. Averaged brain activation patterns across subjects in the circled representative time points are shown for each brain state. (b)  $B_T$  averaged over all the time points in each brain state. (c) Two-dimensional view of task trajectories with the embedding points. Trajectories are colored by each task and data points are colored by the brain states as in (a)

quantify the distribution of states during rest, we assigned each resting-state time point to one of the four previously identified brain states based on the brain state of the nearest task time point. As with the task data, we next calculated the brain state dwell time distribution across the entire resting-state scan (Figure 7b). A non-uniform dwell-time distribution is discovered, with fixation and transition states having a higher proportion of time points than the cognitive states ( $\chi^2 = 205, df = 3, p < .001$ ). Except for the lower-level cognition and the transition states in the social task (which have very few time points to robustly calculate entropy, see Figure 7c), all states exhibit higher entropy in the resting state than during a given task.

In Figure S2, we plot the extension of the WM task. The 2-back and 0-back task blocks go to the correct higher-level cognition or lower-level cognition state respectively, while the fixation and cue time frames are also located in the correct brain states. The correlation between the extended coordinates and the coordinates from the

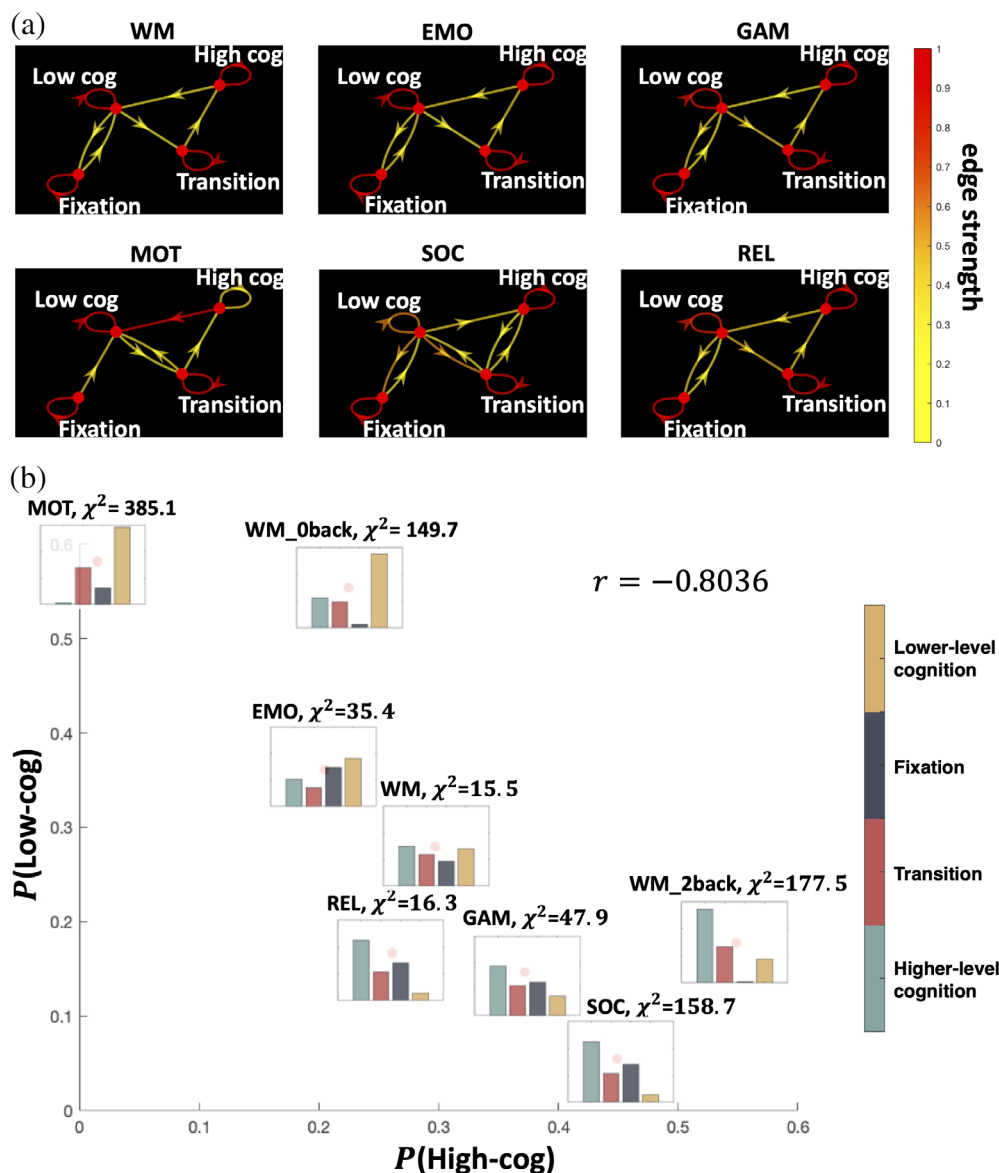
original embedding was highly significant ( $r = .939, p < .001$ ). Holding out the other tasks produced similar results as the WM task.

### 3.5 | Replication of embedding

Notably, we replicated the dimensionality reduction result using participants from the CNP dataset. A similar low-dimensional structure, brain states were found, verifying the robustness of the observed embeddings (Figure 8). Moreover, the same task scans from the schizophrenia cohorts were also embedded separately and found to be similar to the embedding from the HCP dataset and healthy control cohorts in the CNP dataset (Figure S8). This laid foundation for the downstream brain dynamics analysis (resting-state brain dynamics) that would be based on brain states as similar brain states could be identified in both groups.

**FIGURE 6** Brain state dynamics differ between tasks.

(a) Brain state dynamics visualized as the Markov chain. Transition probability is visualized by the color of the directed edges. Edges with transition probability less than 0.03 are omitted for clarity. (b) Stationary distribution probability visualized for each task and positioned by the proportion of higher-level cognition and lower-level cognition brain states. Chi-square test result against the uniform distribution is also shown



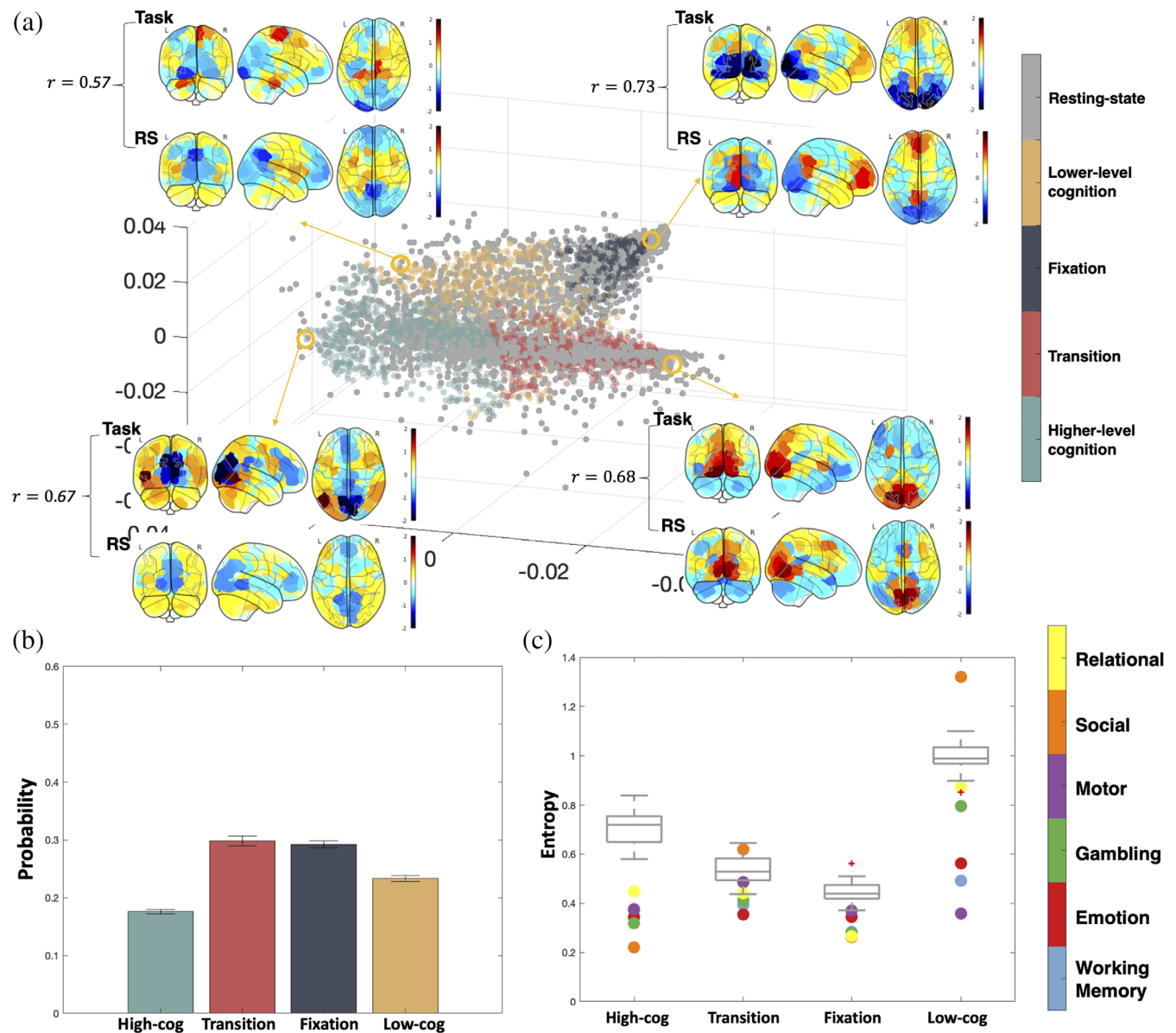
## 4 | DISCUSSION

Using a recently validated manifold learning framework, named 2-step Diffusion Maps—2sDM (Gao et al., 2019), we demonstrate that fMRI data from different tasks span the same low-dimensional embedding (i.e., brain states). In other words, moment-to-moment dynamics from any of these tasks group into the same small number of representative patterns that are hidden from direct observation. The embedding maintained proper temporal progression of the tasks, revealing brain states and temporal dynamics of changes in network integration. Further, we demonstrate that resting-state data project onto the same task embedding using a specially designed out-of-sample-extension method, indicating similar brain states are present. Finally, we validate this embedding using an independent dataset.

Several other publications have organized the temporal dynamics of the brain into a low dimension space or into distinct brain states (Allen et al., 2014; Saggari et al., 2018; Vidaurre, Smith, & Woolrich, 2017) using

data from resting-state or a single task to construct the embedding (Gallego et al., 2017; Shine et al., 2019). Together, these works suggest that a low-dimensional structure exists; however, it is unclear how these structures adapt to diverse cognitive loads. By projecting a rich repertoire of task data into a single manifold, we show that, across different tasks, parts of the embedding (i.e., brain states) are well characterized by the extent to which the brain is integrated. Overall, the discrete states and association with complex network measures suggest that our embedding finds an intrinsic, latent structure of brain dynamics.

These results are in line with the theory that the brain is able to reconfigure its large-scale organization dynamically either between different cognitive tasks or within resting-state (Cohen & D'Esposito, 2016; Shine et al., 2016). Further, they emphasize that this reconfiguration is shared across different cognitive loads and, importantly, resting-state. In other words, the same highly integrated state that characterizes a cognitively demanding task, such as a 2-back WM task, can be observed during resting-states and less

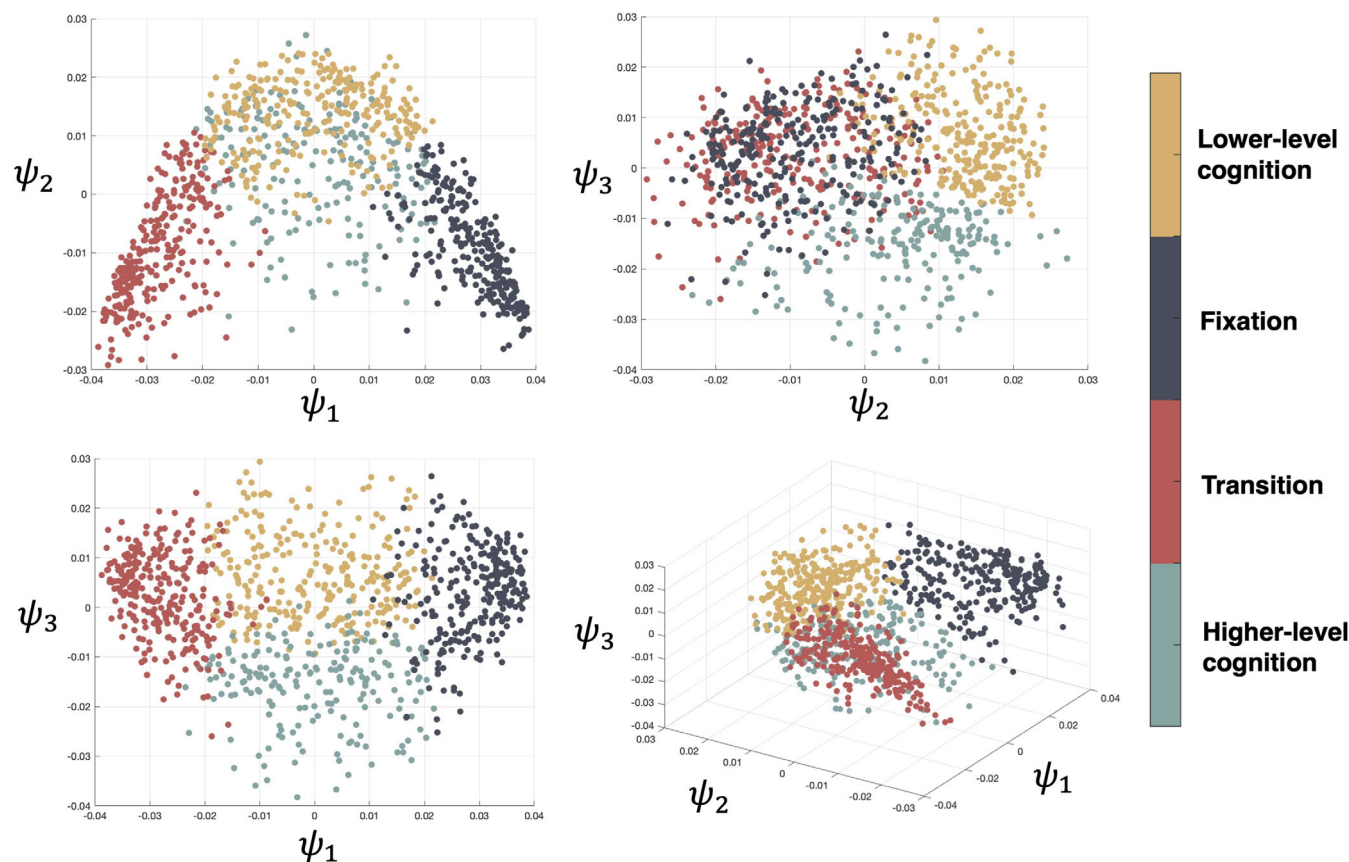


**FIGURE 7** Resting-state extended onto the task manifold. (a) Representative task activation patterns of each state and the neighboring resting-state activation pattern are visualized. Correlation of the activation between task and rest is calculated with higher correlation representing more accurate out-of-sample extension. (b) Stationary probability distribution of the four brain states during resting state. (c) Entropy of each brain state's transition probability in different tasks. Dots are colored by tasks they represent, and the gray box plot shows the entropy values of resting state with BrainSync (see Section 2) referenced to different individuals

cognitively demanding tasks, just with less frequency. These states can also be viewed from a dynamic system perspective (Taghia et al., 2018). As clustering based on the eigenvectors of the normalized graph Laplacian has been used to find meta-stable state in the stochastic dynamical systems (Huisinga, Best, Roitzsch, Schütte, & Cordes, 1999), the four brain states defined from the task scan can also be viewed as four different metastable states. Further, the temporal trajectories can separate different portions of tasks based on cognitive demand, suggesting a potential utility of the embedding for other downstream analyses of brain dynamics.

In line with this, the dynamics between states, rather than within brain states themselves, appear to be the key distinguishing factor

between task and rest. In support of this, how the brain transitions between different states is dependent on the task being performed and is less predictable in resting-state compared to tasks. Executing a task limits the transitions between states; while, during resting-state, the brain can more liberally traverse through different states. Though speculative, these results offer an explanation as to why task connectivity data is better at identifying individuals and subsequent predicting behaviors than resting-state connectivity data (Finn et al., 2017; Greene, Gao, Scheinost, & Constable, 2018). Together, while the resting state may exhibit similar states as observed during task, the temporal dynamics of switching states are less predictable in resting state compared to task.



**FIGURE 8** 2sDM embedding and *k*-means clustering result of CNP dataset. CNP, Consortium for Neuropsychiatric Phenomics; 2sDM, 2-step Diffusion Maps

Previous work demonstrates that brain networks fluctuate between states of low and high global integration during tasks as characterized by the participation coefficient ( $B_T$ ) from sliding-window functional connectivity. Tasks requiring higher cognitive loads, such as the 2-back condition in the WM task, exhibit greater integration while less cognitive load, such as the motor task, exhibits lower integration (Shine et al., 2016). A key drawback of these results is that they rely on two intermediate steps (e.g., the method used to construct dynamic functional connectivity and topological metrics to study), rather than the learned features from unsupervised methods. Together, our results suggest that the task embedding reveals latent information about changes in network topology without the need for handcrafted features. For example, each task can be effectively characterized from the proportion of time spent in lower-level and higher-level cognition states creating a similar ordering of task (see Figure 6b) as in (Shine et al., 2016).

While resting-state fMRI is a powerful tool to map the functional organization of the brain, inherent limitations exist. Resting-state is often conceptualized as a single task state. Though emerging data, including our results, suggest that resting-state is not one single, monolithic state, but rather a collection of multiple states associated with different cognitive loads that also appear during tasks (Cornblath et al., 2020; van der Meer, Breakspear, Chang, Sonkusare, & Cocchi, 2020; Vidaurre et al., 2017). In this paper, we find that

resting-state is dominated by both the fixation and transition state. This aligns well with the previous research that also illustrates this bistable dynamic (Cornblath et al., 2020; van der Meer et al., 2020; Vidaurre et al., 2017). Moreover, while the majority of resting-state time points cluster into a single part of the manifold (such as the fixation blocks, which putatively are the most like “rest”), nearly a third of the time points more closely match cognitive states. Perhaps, more importantly different groups may have differences in “performing” rest (Buckner, Krienen, & Yeo, 2013). How best to interpret changes in resting-state connectivity in the presence of group differences in dynamics is still an open question.

A key strength of our embedding framework is its data-driven nature. We demonstrated that the embedding coordinates could reveal topological information originally found using dynamic functional connectivity methods (Shine et al., 2016). This brain topology was found without specifying common modeling choices in dynamic functional connectivity, such as how to model the functional connectivity (i.e., statistical interdependence of signals) between brain regions, an underlying graph/network, or even information about task stimuli (e.g., block lengths). As a multitude of methodological choices have been proposed to analyses (Calhoun, Miller, Pearlson, & Adali, 2014; Hutchison et al., 2013; e.g., ways of estimating connectivity [Allen et al., 2014; Chang & Glover, 2010; Shine et al., 2015], constructing a weighted or unweighted graph [Rubinov & Sporns, 2010],



specific graph theory measures [Honey, Kotter, Breakspear, & Sporns, 2007; Meunier, Lambiotte, & Bullmore, 2010; Shine et al., 2016; Sizemore & Bassett, 2018]), our embedding framework provides an end-to-end, data-driven approach without the need for modeling choices to investigate brain dynamics. More generally, hand-crafted features are being substituted by more automatic feature learning-based nonlinear methods such as deep learning and nonlinear embedding methods (Hamilton, Ying, & Leskovec, 2017). Our results show a specific scenario in which “let the data speak for itself” is an achievable option for modeling fMRI data.

A limitation of this work is that the embedding can only “look under the light.” That is to say that, while a rich amount of task data was needed to create the embedding, we could not include every possible task in creating the embedding. Indeed, it is highly likely that many more than four brain states exist and that we have not detected every single one. A finer grade delineation of states, probably through further advancement in non-linear embedding methods, is a needed future direction of work. Moreover, although here brain states are defined based on the  $k$ -means clustering result, it does not rule out other ways to define brain states. For example, at each time point, the brain can also be modeled as being at different states with distinct probabilities (Vidaurre et al., 2017), which can be achieved by a fuzzy-clustering algorithm. Moreover, the brain state can also be characterized by the temporal trajectory, Trajectory clustering technique can be used to cluster trajectories into trajectory-based brain states, which takes account the temporal information of the embedding (Lee, Han, & Whang, 2007). Our  $k$ -means clustering approach to defining brain state is only one of the ways to summarize information of the embedding and serves as a proof-of-concept that our embedding contains information that is relevant to brain dynamics. Nevertheless, the observed task embedding was similar across two different input datasets with different tasks, suggesting that embedding is general to factors such as scanner, task, processing, and sample size.

One of the assumptions of 2sDM is that the time frames from all individuals are temporally aligned so that a group-average embedding of the time frames can be obtained. However, this does not rule out the applicability of the task scans that has different task block lengths/orders across individuals (e.g., language task in the HCP dataset) or the resting-state scans, which we have demonstrated in the paper by applying BrainSync. Thus, task scans with distinct block lengths/orders can also be embedded with 2sDM by applying BrainSync first. It is worth noting that as BrainSync requires a specific individual chosen as the reference, by aligning all the other individuals to the same selected individual, the group-average embedding then will approximate a cleaner temporal embedding of the selected individual, which can be used to investigate individual-level dynamics.

Finally, 2sDM relies on diffusion maps to perform the embedding. Other nonlinear embedding methods exist, such as t-Distributed Stochastic Neighbor Embedding (t-SNE) and Potential of Heat-diffusion for Affinity-based Trajectory Embedding (PHATE). Future work will be to investigate the potential of these methods in embed timeseries from a rich repertoire of tasks onto a single low-dimensional manifold in using a similar two step framework. While nonlinear embedding

methods are better able to reveal manifold structures, the mappings between the raw neural space and the low-dimensional coordinates are harder to interpret. Future work will also address in more details that how the neural spaces is interrelated with the 2sDM coordinates.

The ability to use data-driven methods to clearly identify a low-dimensional space of brain dynamics, regardless of how the brain is engaged during imaging, indicates that these brain dynamics are robust and reliable across conditions in addition to being unique. Together, these advances suggest that analysis of individual fMRI data from multiple cognitive tasks in a low-dimensional space is possible, and indeed, desirable.

## ACKNOWLEDGMENTS

This work was funded by NIH MH121095. Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657 funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University) and the Consortium for Neuropsychiatric Phenomics (NIH Roadmap for Medical Research grants UL1-DE019580, RL1MH083268, RL1MH083269, RL1DA024853, RL1MH083270, RL1LM009833, PL1MH083271, and PL1NS062410).

## CONFLICT OF INTERESTS

The authors declare no competing financial interests.

## DATA AVAILABILITY STATEMENT

The HCP data used in this study are publicly available from the ConnectomeDB database (<https://db.humanconnectome.org>). The CNP data used in this study are publicly available from OpenNeuro.org (<https://openneuro.org/datasets/ds000030>). MATLAB scripts to run the 2sDM analyses can be found at (<https://github.com/carricky/2sDM>). Biolmage Suite tools used for analysis can be accessed at (<https://bioimagesuiteweb.github.io/>).

## ORCID

Dustin Scheinost  <https://orcid.org/0000-0002-6301-1167>

## REFERENCES

- Ahrens, M. B., Li, J. M., Orger, M. B., Robson, D. N., Schier, A. F., Engert, F., & Portugues, R. (2012). Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature*, *485*, 471–477.
- Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, *24*, 663–676.
- Buckner, R. L., Krienen, F. M., & Yeo, B. T. T. (2013). Opportunities and limitations of intrinsic functional connectivity MRI. *Nature Neuroscience*, *6*, 832–837.
- Calhoun, V. D., Miller, R., Pearlson, G., & Adali, T. (2014). The chronnectome: Time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron*, *84*, 262–274.
- Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*, 1–27.
- Chang, C., & Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage*, *50*, 81–98.



- Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., & Shenoy, K. V. (2012). Neural population dynamics during reaching. *Nature*, *487*, 51–56.
- Cohen, J. R., & D'Esposito, M. (2016). The segregation and integration of distinct brain networks and their relationship to cognition. *The Journal of Neuroscience*, *36*, 12083–12094.
- Coifman, R. R. & Lafon, S. (2006). Diffusion maps. pp. 5–30.
- Cornblath, E. J., Ashourvan, A., Kim, J. Z., Betzel, R. F., Ciric, R., Adebimpe, A., ... Moore, T. M. (2020). Temporal sequences of brain activity at rest are constrained by white matter structure and modulated by cognitive demands. *Communications Biology*, *3*, 1–12.
- Cunningham, J. P., & Byron, M. Y. (2014). Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, *17*, 1500.
- Finn, E. S., Scheinost, D., Finn, D. M., Shen, X., Papademetris, X., & Constable, R. T. (2017). Can brain state be manipulated to emphasize individual differences in functional connectivity? *NeuroImage*, *160*, 140–151.
- Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., ... Constable, R. T. (2015). Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. *Nature Neuroscience*. Nature Publishing Group, *18*, 1664–1671.
- Fowlkes, C., Belongie, S., Chung, F., & Malik, J. (2004). Spectral grouping using the Nystrom method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *26*, 214–225.
- Gallego, J. A., Perich, M. G., Miller, L. E., & Solla, S. A. (2017). Neural manifolds for the control of movement. *Neuron*, *94*, 978–984.
- Gao, P., & Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. *Current Opinion in Neurobiology*, *32*, 148–155.
- Gao, S., Mishne, G., & Scheinost, D. (2019). A Hierarchical Manifold Learning Framework for High-Dimensional Neuroimaging Data. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer International Publishing, pp. 631–643.
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., ... Consortium, W.-M. H. (2013). The minimal preprocessing pipelines for the human connectome project. *NeuroImage*, *80*, 105–124.
- Greene, A. S., Gao, S., Scheinost, D., & Constable, R. T. (2018). Task-induced brain state manipulation improves prediction of individual traits. *Nature Communications*, *9*, 2807.
- Hamilton, W.L., Ying, R., & Leskovec, J. (2017). Representation learning on graphs: Methods and applications. arXiv preprint arXiv:1709.05584.
- Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., ... Ramadge, P. J. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*, *72*, 404–416.
- Honey, C. J., Kotter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 10240–10245.
- Huisinga, W., Best, C., Roitzsch, R., Schütte, C., & Cordes, F. (1999). From simulation data to conformational ensembles: Structure and dynamics-based methods. *Journal of Computational Chemistry*, *20*, 1760–1774.
- Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., ... Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage*, *80*, 360–378.
- Jorgenson, L. A., Newsome, W., Anderson, D. J., Bargmann, C. I., Brown, E. N., Deisseroth, K., ... Wingfield, J. C. (2015). The BRAIN initiative: Developing technology to catalyze neuroscience discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1668), 20140164.
- Joshi, A., Scheinost, D., Okuda, H., Belhachemi, D., Murphy, I., Staib, L. H., & Papademetris, X. (2011). Unified framework for development, deployment and robust testing of neuroimaging algorithms. *Neuroinformatics*, *9*, 69–84.
- Joshi, A. A., Chong, M., Li, J., Choi, S., & Leahy, R. M. (2018). Are you thinking what I'm thinking? Synchronization of resting fMRI time-series across subjects. *NeuroImage*, *172*, 740–752.
- Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, *32*, 922–923.
- Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C. E., Kepecs, A., Mainen, Z. F., ... Machens, C. K. (2016). Demixed principal component analysis of neural population data. *Elife*, *5*, e10989.
- Lee, J. -G., Han, J., & Whang, K. -Y. (2007). Trajectory clustering: a partition-and-group framework. Proceedings of the 2007 ACM SIGMOD international conference on Management of data, pp. 593–604.
- Lutkenhoff, E. S., Rosenberg, M., Chiang, J., Zhang, K., Pickard, J. D., Owen, A. M., & Monti, M. M. (2014). Optimized brain extraction for pathological brains (optiBET). *PLoS One*, *9*, e115551.
- Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., ... Petrides, M. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, *113*, 12574–12579.
- Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience*, *4*, 200.
- Mishne, G., Talmon, R., Meir, R., Schiller, J., Lavzin, M., Dubin, U., & Coifman, R. R. (2016). Hierarchical coupled-geometry analysis for neuronal structure and activity pattern discovery. *IEEE Journal of Selected Topics in Signal Processing*, *10*, 1238–1253.
- Monti, R. P., Lorenz, R., Hellyer, P., Leech, R., Anagnostopoulos, C., & Montana, G. (2017). Decoding time-varying functional connectivity networks via linear graph embedding methods. *Frontiers in Computational Neuroscience*, *11*, 14.
- Nenning, K. H., Xu, T., Schwartz, E., Arroyo, J., Woehrer, A., Franco, A. R., ... Langs, G. (2020). Joint embedding: A scalable alignment to compare individuals in a connectivity space. *NeuroImage*, *222*, 117232.
- Poldrack, R. A., Congdon, E., Triplett, W., Gorgolewski, K. J., Karlsgodt, K. H., Mumford, J. A., ... Bilder, R. M. (2016). A phenome-wide examination of neural and cognitive function. *Scientific Data*, *3*, 160110.
- Preti, M. G., Bolton, T. A., & Van De Ville, D. (2017). The dynamic functional connectome: State-of-the-art and perspectives. *NeuroImage*, *160*, 41–54.
- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*, 1059–1069.
- Saggar, M., Sporns, O., Gonzalez-Castillo, J., Bandettini, P. A., Carlsson, G., Glover, G., & Reiss, A. L. (2018). Towards a new approach to reveal dynamical organization of the brain using topological data analysis. *Nature Communications*, *9*, 1–14.
- Santhanam, G., Yu, B. M., Gilja, V., Ryu, S. I., Afshar, A., Sahani, M., & Shenoy, K. V. (2009). Factor-analysis methods for higher-performance neural prostheses. *Journal of Neurophysiology*, *102*, 1315–1330.
- Scheinost, D., Kwon, S. H., Lacadie, C., Vohr, B. R., Schneider, K. C., Papademetris, X., ... Ment, L. R. (2017). Alterations in anatomical covariance in the prematurely born. *Cerebral Cortex*, *27*, 534–543.
- Shen, X., Tokoglu, F., Papademetris, X., & Constable, R. T. (2013). Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *NeuroImage*, *82*, 403–415.
- Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., ... Poldrack, R. A. (2016). The dynamics of functional brain networks: Integrated network states during cognitive task performance. *Neuron*, *92*, 544–554.
- Shine, J. M., Breakspear, M., Bell, P. T., Ehgoetz Martens, K. A., Shine, R., Koyejo, O., ... Poldrack, R. A. (2019). Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nature Neuroscience*, *22*, 289–296.

- Shine, J. M., Koyejo, O., Bell, P. T., Gorgolewski, K. J., Gilat, M., & Poldrack, R. A. (2015). Estimation of dynamic functional connectivity using multiplication of temporal derivatives. *NeuroImage*, *122*, 399–407.
- Sizemore, A. E., & Bassett, D. S. (2018). Dynamic graph metrics: Tutorial, toolbox, and tale. *NeuroImage*, *180*, 417–427.
- Taghia, J., Cai, W., Ryali, S., Kochalka, J., Nicholas, J., Chen, T., & Menon, V. (2018). Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. *Nature Communications*, *9*, 1–19.
- van der Meer, J. N., Breakspear, M., Chang, L. J., Sonkusare, S., & Cocchi, L. (2020). Movie viewing elicits rich and reliable brain state dynamics. *Nature Communications*, *11*, 1–14.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., & Consortium, W.-M.H.C.P. (2013). The WU-Minn human connectome project: An overview. *NeuroImage*, *80*, 62–79.
- Vidaurre, D., Smith, S. M., & Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, *114*, 12827–12832.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Gao, S., Mishne, G., & Scheinost, D. (2021). Nonlinear manifold learning in functional magnetic resonance imaging uncovers a low-dimensional space of brain dynamics. *Human Brain Mapping*, *42*(14), 4510–4524. <https://doi.org/10.1002/hbm.25561>